(2)

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
51 HOVEY ROAD, PENSACOLA, FL 32508-1046

NAMRL SPECIAL REPORT 93-3

A COMPARISON OF TWO HARDWARE IMPLEMENTATIONS OF A CREW SELECTION SYSTEM

T. Nontasak and K. T. Helton

Approved for public release; distribution unlimited.

Reviewed and approved 8 Oct 93

A. J. MATECZUN, CAPT, MC USN
Commanding Officer

DEPARTMENT OF THE NAVY
AEROSPACE MED RESEARCH LABORATORY

This research was sponsored by the Naval Sea Systems Command (Code PMS 377) through the Naval Command Control Ocean Surveillance Center and supported by the Naval Coastal Systems Center under Work Unit Number CE02. Additional support was provided by the Naval Medical Research and Development Command under work unit No. 63706N M0096.002-M00960.01.

The views expressed in this report are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

Trade names of materials and/or products of commercial or nongovernment organizations are cited as needed for precision. These citations do not constitute official endorsement or approval of the use of such materials and/or products.

Volunteer subjects were recruited, evaluated and employed in accordance with the procedures specified in the Department of Defense Directive 3216.2 and the Secretary of the Navy Instruction 3900.39 series. These instructions are based upon voluntary informed consent and meet or exceed the provisions of prevailing national and international guidelines.

ABSTRACT

A number of disadvantages existed within the Apple IIe version of the Landing Craft Air Cushion (LCAC) vehicle Crew Selection System (LCSS). These disadvantages included slow processing speed, memory constraints, and cumbersome test administration. The system was also less conducive to future test development. The LCSS was upgraded to an IBM PC compatible Zenith-248 system. A comparison of 48 single and composite subtest scores between like measures of the two systems yielded significant positive correlations ($p < .0001$). The majority of the correlations ranged from .60 to .89. A small percentage of the associations were less robust. Overall, however, the Apple IIe and Zenith 248 versions of the LCSS proved to be comparable testing systems.

Acknowledgments

The authors gratefully acknowledge P.D. Collyer, LT. M.D. Reddix, and K. S. Mayer for reviewing this manuscript. We also wish to acknowledge the initial development work of LCDR G.D. Gibb and G.R. Griffin, technical information and programming efforts of P.D. Collyer, and test administration by A. Thomas, HM1 S.D. Eagles, and S.P. Starling. We are especially grateful for the support given by Dr. J.O. de Lorge, LCDR D.L. Dolgin, LCDR D.J. Blower, LT D.R. Street, Jr., Dr. H.D. Delaney, I.K. Vogel, and A. D. Chapman. The support of the Naval Sea Systems Command's Amphibious Warfare Program, particularly by Dr. L.E. Hufford of NCCOSC/NRAD 44211, is greatly appreciated. We also thank all the subjects who participated in the study.

INTRODUCTION

Advances in computer technology offer opportunities to improve computer-based crew selection instruments previously developed by the Naval Aerospace Medical Research Laboratory (NAMRL). The Navy's Landing Craft Air Cushion (LCAC) vehicle Crew Selection System (LCSS) was recently upgraded from an Apple IIe-based system to a Zenith 248-based system. This paper presents the comparison of general hardware features and the results of analyses conducted to determine the testing equivalence of the two computer-based systems.

In 1988, NAMRL developed the Apple IIe-based LCSS to reduce high attrition rates (35-40%) among LCAC crew trainees (Eakin, 1990). As training costs escalated and mission necessities demanded more LCAC vehicles and crew, identifying successful trainee candidates became crucial. No systematic LCAC personnel selection system existed. The development of the selection system was based on extensive interviews with LCAC fleet craftmasters, training personnel, and program managers. The information provided in these interviews led to the utilization of the NAMRL in-house aviation selection test battery as an experimental test battery for LCAC crew and candidates. This battery became the LCSS. The battery included tests to measure psychomotor skills, reaction time, time-sharing abilities, biographical characteristics, and personality traits. A series of concurrent and predictive validation studies demonstrated that the LCSS was predictive of performance in the first phase of LCAC training (Dolgin & Nontasak, 1990; Dolgin, Street, Nontasak, Blower, & Travis, 1992; Nontasak & Dolgin, 1990; Nontasak, Dolgin, & Griffin, 1989). The Apple IIe-based LCSS was operationally implemented in 1990 and replaced by the IBM PC compatible Zenith 248-based version in 1992. Since the implementation of the LCSS, trainees' attrition rates have dropped below 10% (Nontasak, Dolgin, Helton, Street, & Blower, 1993).

Although proven a useful tool for the LCSS, the Apple IIe contained several drawbacks compared to contemporary computing machines: slow processing speed, memory constraints, and cumbersome test administration. As noted by Eamon and Butler (1987), the Apple IIe was underpowered, poorly designed, and generally obsolete in terms of hardware. The IBM-compatible Zenith 248 was selected to replace the Apple IIe because of its faster speed, larger memory capacity, ease in test administration, and flexibility in software development. More detailed descriptions of the two systems can be found in the Apple IIe owner's manual (1983) and the Zenith 248 owner's manual (1986).

The literature in research instrumentation and techniques indicated concerns regarding the use of different types of hardware to evaluate the same human performance measures. Specifically, the problem was whether data obtained from one machine were comparable to data obtained from another (Graves and Bradley, 1988; Kane & Kay, 1992; Segalowitz and Graves, 1990). We investigated the extent to which test performance data obtained from the Apple IIe-based LCSS were comparable to data obtained from the Zenith 248-based system.

METHOD

SUBJECTS

The subjects consisted of 106 male LCAC-trainee candidates ranging in age from 22 to 46 years (M = 31.98, SD = 4.12). They were high school or GED graduates and had extensive shipboard experience. The candidates were previously screened and recommended by their respective commands. As part of the Naval Bureau of Personnel prescreening for LCAC training, they were required to obtain a minimum score of 240 on selected Armed Services Vocational Aptitude Battery subtests. All subjects participated in the study as part of the selection process between May 1990 and October 1992.

INSTRUMENTS

Hardware. Hardware components of the Apple IIe system are shown in Fig. 1; Zenith-248 system components are illustrated in Fig. 2. Although some of the components of the two systems are identical (joysticks, rudder pedals, headphones, and cassette tape player/recorder), several critical components differ. Both systems use voice synthesizers, however, the Zenith uses a voice synthesizer incorporated within the LABPAC multifunction board. The Apple IIe system requires the tests to be administered via floppy disks, which test administrators change between tests. All tests with the Zenith-248 are administered directly from a single 20-MB hard drive. In addition, the Zenith-248 system is enclosed in a console, providing better human factors engineering quality for the subjects. The console also reduces extraneous noise and cathode-ray-tube glare and provides a more durable testing station.

Figure 3 demonstrates the differences in the location of the joysticks for the two systems. The center joystick (stick) for the Apple IIe-based system was mounted in the center on the forward edge of a standard straight-backed metal chair. The joystick on the left (throttle) was mounted on the left edge of the chair. For the Zenith system, the stick was mounted on the forward edge of the testing console at a center position. The throttle was located on the left side of the testing console.

Software. The test programs for both systems were written in GW BASIC programming language. These programs ran several performance-based tests measuring eye-hand-foot coordination skills. The two psychomotor tests making up the LCSS include the Psychomotor and Dichotic Listening Test (PMT/DLT) and the Compensatory Tracking/Digit Cancellation (CT/DC) test. A brief description of the PMT/DLT and the CT/DC is provided below. A detailed description can be found elsewhere (Helton, Nontasak, & Dolgin, 1992).

The PMT/DLT, composed of seven subtests, was designed to measure psychomotor abilities and divided attention. The PMT portion of the test required the subject to maintain cursors, controlled by joysticks and rudder pedals, on appropriate targets. The DLT involved the presentation of a different set of letters and numbers to each ear simultaneously. The computer specified which ear to pay attention to, and the subject keyed in only the numbers heard in that ear. The subject began with the simplest task of maintaining only one cursor on target. Later, the subject continued with the more difficult task of maintaining three cursors on target and performing the DLT simultaneously.

The CT/DC test also measured psychomotor and time-sharing abilities. In this test, subjects were required to keep only one cursor on target. However, a computer-programmed forcing function made this task difficult by requiring subjects to make continuous counterbalance movements to keep the cursor on target. The digit cancellation task required the subject to key in numbers as they appeared on the screen. Both reaction time and accuracy were recorded. The final task required the subject to perform both tasks simultaneously.

PROCEDURE

All subjects were tested first on the Apple IIe systems used for the actual selection of the candidates. After completing the selection tests, they were asked to repeat the two psychomotor tests on the Zenith-248 systems. The time between testing on the two systems generally ranged from 15 min to 3 h depending on subject preference and test station availability. On both systems, subjects performed the PMT/DLT tests followed by the CT/DC tests.

Figure 1. The Apple IIe System.

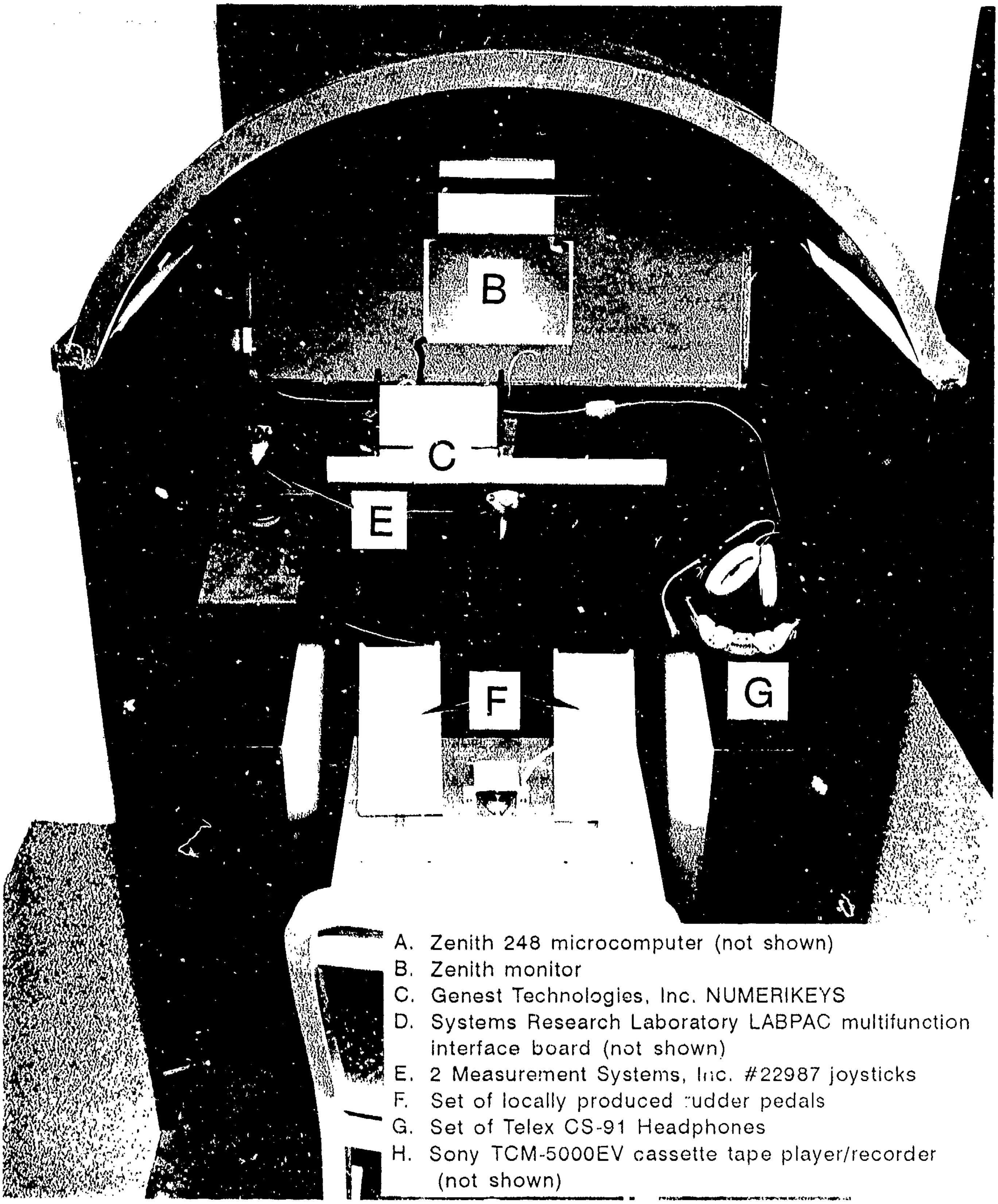
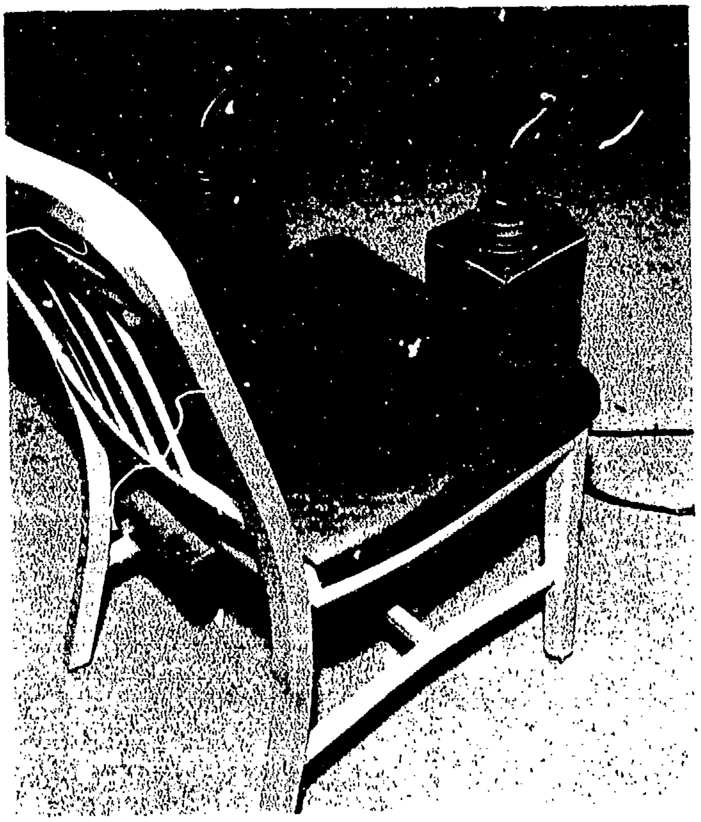


Figure 2. The Zenith System.



Apple IIe system

Zenith 248 system

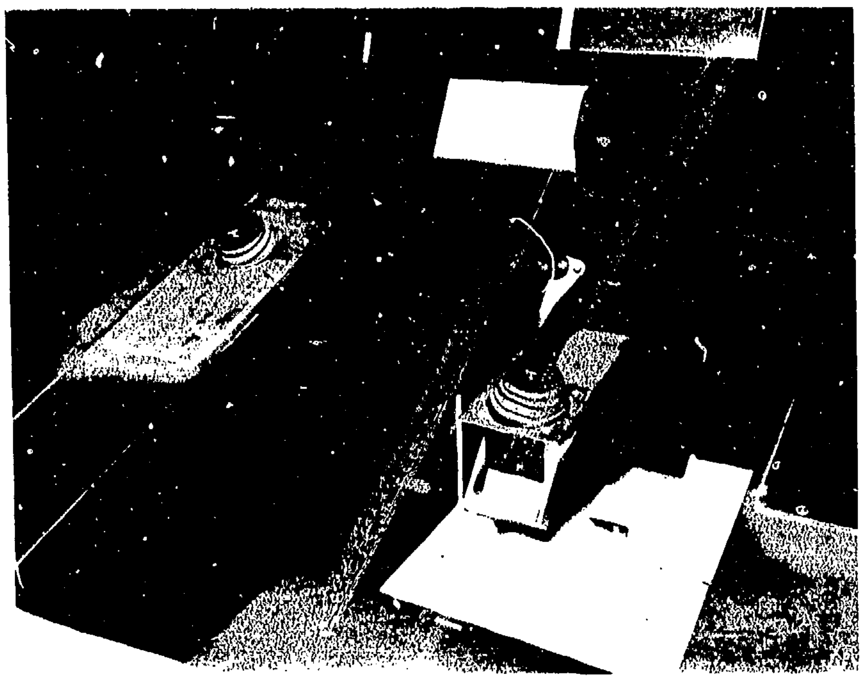


Figure 3. Joysticks location comparison.

RESULTS

Test scores on both systems were subjected to a Pearson r correlational analysis to determine the comparability of the two systems. The results of these analyses are shown in Tables 1-4.

Table 1 shows the relationship between the DLT tasks on the two systems. For all four DLT subtasks and for the DLT composite score the correlations were significant at the .0001 level.

Table 1. Apple-Zenith Correlation Coefficients of DLT Scores ($N = 106$)

DLT Correct Scores	
Test	r
DLT - single task only ($N = 105$)	$.80^{*}$
DLT - with stick task	$.76^{*}$
DLT - with stick & rudder	$.82^{*}$
DLT - with stick, rudder, & throttle	$.71^{*}$
DLT composite	$.87^{*}$

$^{*}p < .0001$

The correlations for the PMT subtest scores are shown in Table 2. All analyses, including the analyses of the two composite scores (PMT composite and PMT-DLT composite), were significant at the .0001 level.

Table 2. Apple-Zenith Correlation Coefficients of PMT Error Scores ($N = 106$)

PMT Error Scores	
Subtests	r
Stick only	.73*
Stick & DLT	.76*
Stick & rudder	.67*
Stick, rudder & DLT	.83*
Stick, rudder, & throttle:	
stick & rudder (1st session)	.75*
stick & throttle (1st session)	.74*
stick, rudder, & throttle (1st session)	.74*
stick & rudder (2nd session)	.76*
stick & throttle (2nd session)	.80*
stick, rudder, & throttle (2nd session)	.76*
stick, rudder, & throttle (1st session + 2nd session)	.80*
Stick, rudder, throttle, & DLT:	
stick & rudder	.75*
stick & throttle	.68*
stick, rudder, & throttle	.72*
PMT/DLT Composite scores	
PMT composite	.82*
PMT/DLT composite z score ($N = 105$)	.89*

* $p < .0001$

Table 3 illustrates the relationship between both the Apple and Zenith scores for the CT single tracking task. For all seven trials and the average of the last three trials, performance was equivalent across the two systems ($p < .0001$).

Table 3. Apple-Zenith Correlation Coefficients of Single Tracking Error Scores (N = 106)

Compensatory Tracking Raw Error Scores	
Trial	r
1	.47*
2	.60*
3	.55*
4	.52*
5	.47*
6	.48*
7	.46*
Average of last 3 trials	.50*

* $p < .0001$

The DC reaction time (RT) and accuracy scores were also significantly correlated ($p < .0001$). The r values for correct RT, overall RT, and correctly canceled digits were .85, .85, and .88, respectively.

The relationship between scores for the CT/DC dual task are shown in Table 4. All associations were significant.

Table 4. Apple-Zenith Correlation Coefficients of CT/DC Dual Task Scores (N = 106)

Dual Task Raw Scores	
Subtest scores	r
Correct RT - trial 1	.72*
Correct RT - trial 2	.73*
Correct RT - trial 3 (N = 105)	.75*
Overall RT (average of 3 trials)	.81*
Overall RT (average of last 2 trials)	.74*
Correctly canceled digits - trial 1	.79*
Correctly canceled digits - trial 2 (N = 104)	.81*
Correctly canceled digits - trial 3 (N = 105)	.86*
Tracking error - trial 1	.57*
Tracking error - trial 2	.63*
Tracking error - trial 3	.58*
Average of digits correctly canceled (3 trials)	.85*
Average of digits correctly canceled (last 2 trials) (N = 104)	.85*
Average of tracking error (3 trials)	.65*
Average of tracking error (last 2 trials)	.64*
Compensatory Tracking/Digit Cancellation composite score	.83*

* $p < .0001$

DISCUSSION AND CONCLUSIONS

Our findings revealed high correlations between the same test performance measures on the Apple IIe-based and the Zenith 248-based LCSS systems. Of the 48 single and composite subtests, 39 (81%) exhibited correlation coefficients of .60 to .89 ($p < .0001$). The average correlation was $r = .72$. These robust associations appear to demonstrate the two systems' comparability. Some correlations were more moderate, but all were significant at the .0001 level. For example, the r for the CT task average was .50 ($p < .0001$). Overall, the correlational analyses demonstrated that the two systems were comparable.

Practice or learning effects have always been a concern when repeated measures of the same tests are taken. Because the subjects took the Apple version of the LCSS followed by the Zenith version, these effects may have contributed to the high correlations obtained in this study. However, previous research by H.D. Delaney (personal communication, July 23, 1991) comparing naval aviator PMT/DLT and CT/DC performance on an Apple-to-Zenith test sequence and a Zenith-to-Apple test sequence indicated that no consistent practice effects existed. We feel that any practice or learning effect could have also been nullified

by the difficulty associated with the faster speed of the Zenith system. Although counterbalancing would have provided an answer to this issue, it would have been impractical. Administering the Zenith version first could have fatigued the subjects and thereby impaired their performance on the Apple selection tests.

Transitioning the LCSS from the Apple IIe-based system to the Zenith 248-based system yielded several advantages. First, the Zenith systems provided faster processing speeds. The Apple system tended to overload or respond slowly when several controllers were functional at the same time. The Zenith's processing speed allowed more performance sampling per millisecond, thereby providing a more accurate depiction of a subject's ability. Second, the larger random access memory and disk space in the Zenith systems allowed the tests to be menu-driven and run directly from the hard drive, whereas the Apple IIe system tests were administered via floppy disks, making test administration cumbersome. Thus, administering the tests on the Zenith was facilitated, and test administrator interference was minimized. Finally, the Zenith systems provided more flexibility for future test modifications or additions.

Enclosing the test equipment in a console also provided some advantages. The console provided a more face-valid testing environment because it resembled the LCAC cockpit. Our posttest debriefing sessions with the subjects indicated that they preferred the Zenith partly because it was better constructed to make the user more comfortable (e.g., the controllers were more easily accessible). The console also reduced extraneous distractions (noise), eliminated light glare on the screen, and reduced potential safety hazards (tripping over wires). Finally, the console protected system components and made the test stations more presentable and durable.

REFERENCES

Apple IIe owner's manual. (1983). Cupertino, CA: Apple Computer, Inc.

Dolgin. D.L., and Nontasak, T. (1990). Initial validation of a personnel selection system for landing craft air cushion (LCAC) vehicle operators. *Proceedings of the 12th Symposium Psychology in the Department of Defense* (pp. 245-249). Colorado Springs, CO: USAF Academy Department of Behavioral Sciences and Leadership.

Dolgin, D.L., Street, D.R., Jr., Nontasak, T., Blower, D.J., and Travis, K. (1992). Operational implementation of a validated personnel selection system for landing craft air cushion (LCAC) vehicle operators. *Proceedings of the 13th Symposium Psychology in the Department of Defense* (pp. 66-70). Colorado Springs, CO: USAF Academy Department of Behavioral Sciences and Leadership.

Eakin, D.M. (1990). *Student attrition update* (ltr DME-90-217 of 11 Sep). Panama City, FL: Textron Marine Systems.

Eamon, D.B., & Butler, D.L. (1987). Writing programs on the Apple IIe for the student psychology laboratory: Routines, subroutines, and sources of information. *Behavior Research Methods, Instruments, & Computers, 19* (2), 88-98.

Graves, R., & Bradley, R. (1988). More on millisecond timing and tachistoscope applications for the IBM PC. *Behavior Research Methods, Instruments, & Computers, 20* (4), 408-412.

Helton, K.T., Nontasak, T., & Dolgin, D.L. (1992). *Landing craft air cushion (LCAC) crew selection manual* (NAMRL-1379). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Kane, R.L., & Kay, G.G. (1992). Computerized assessment in neuropsychology: A review of tests and test batteries. *Neuropsychology Review, 3*, pp. 1-117.

Nontasak, T., and Dolgin, D.L. (1990). Differences in time-sharing ability between successful and unsuccessful trainees in the landing craft air cushion vehicle operator training program. *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 949-953). Orlando, FL: Human Factors Society.

Nontasak, T., Dolgin, D.L., and Griffin, R.G. (1989). Performance-based tests, personality attributes, and training outcome among landing craft air cushion (LCAC) vehicle operators. *Proceedings of the Human Factors 33rd Annual Meeting* (pp. 901-904). Denver, CO: Human Factors Society.

Nontasak, T., Dolgin, D.L., Helton, K.T., Street, D.R., & Blower, D.J. (1993). Development and implementation of a Landing Craft Air Cushion vehicle crew selection system. *Proceedings of the 64th Annual Scientific Meeting of the Aerospace Medical Association* (p. A34, #200). Toronto, Canada.

Segalowitz, S.J., & Graves, R.E. (1990). Suitability of the IBM XT, AT, and PS/2 keyboard, mouse, and game port as response devices in reaction time paradigms. *Behavior Research Methods, Instruments, & Computers, 22* (3), 283-289.

Z-248 PC series computers owner's manual. (1986). St. Joseph, MI: Zenith Data Systems Corporation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
October 1993

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE

A Comparison of Two Hardware Implementations of a Crew Selection System

5. FUNDING NUMBERS

63706N
M0096.002-M00960.01

6. AUTHOR(S)

Tatree Nontasak and Kathleen T. Helton

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NAVAEROMEDRSCHLAB
51 HOVEY ROAD
PENSACOLA FL 32508-1046

8. PERFORMING ORGANIZATION REPORT NUMBER

NAMRL Special Report 93-3

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Naval Medical Research and Development Command
National Naval Medical Center, Bldg. 1
Bethesda, MD 20889-5606

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

This work was sponsored by the NAVSEA (Code PMS 377) through the Naval Command Control Ocean Surveillance Center and supported by the Naval Coastal Systems Center (work unit #CE02)

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A number of disadvantages existed within the Apple IIe version of the Landing Craft Air Cushion vehicle Crew Selection System (LCSS). These disadvantages included slow processing speed, memory constraints, and cumbersome test administration. The system was also less conducive to future test development. The LCSS was upgraded to an IBM PC compatible Zenith 248 system. A comparison of 48 single and composite subtest scores between like measures of the two systems yielded significant positive correlations ($p < .0001$). The majority of the correlations ranged from .60 to .89. A small percentage of the associations were less robust. Overall, however, the Apple IIe and Zenith 248 versions of the LCSS proved to be comparable testing systems.

14. SUBJECT TERMS

LCAC, LCSS test battery, Apple IIe, Zenith-248, Computerized tests, Crew selection system

15. NUMBER OF PAGES

13

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std Z39-18
298-102